

# Kritiek op Searle's notie van het formele zelf

---

*Siën J. Rens*

Masterscriptie ter verkrijging van de graad "Master of Arts" in de filosofie

Faculteit der Filosofie, Theologie en Religiewetenschappen  
Radboud Universiteit Nijmegen

Studentennummer: s4204972

Begeleider: L.C. de Bruin

Aantal woorden: 15.175

Datum: 12-08-2020

### **Verklaring van eigen werk**

Hierbij verklaar en verzeker ik, Siën Rens, dat deze scriptie zelfstandig door mij is opgesteld, dat geen andere bronnen en hulpmiddelen zijn gebruikt dan die door mij zijn vermeld en dat de passages in het werk waarvan de woordelijke inhoud of betekenis uit andere werken – ook elektronische media – is genomen door bronvermelding als ontlening kenbaar gemaakt worden.

Plaats: Nijmegen      Datum: 12-08-2020

## **Abstract**

Volgens Searle zijn er een aantal problemen met persoonlijke identiteit. Een van die problemen is of we naast de sequentie van percepties een extra zelf moeten postuleren om onze ervaring van de werkelijkheid te kunnen verklaren. Searle beargumenteert contra-Hume dat dit het geval is. Er moet volgens Searle een formele zelf gepostuleerd worden, omdat we anders niet anders kunnen verklaren waarom (1) de werkelijkheid wordt ervaren als een eenheid en waarom (2) wij het gevoel hebben dat we, op elk moment van de tijd, vrije keuzes kunnen maken. Het formele zelf neemt volgens Searle bewust waar, maakt vrije keuzes en heeft de mogelijkheid om op deze keuzes te kunnen reflecteren. In deze thesis beargumenteer ik dat het niet nodig is om een zelf te postuleren om te begrijpen waarom wij de werkelijkheid als een eenheid ervaren en waarom het voor ons voelt alsof we vrije keuzes maken. De eerste bewering die ik maak is dat de eenheid die wij ervaren niet tot stand komt door een zelf die de verschillende ervaringen van de werkelijkheid aan elkaar verbindt, maar door de werking van hogere orde gedachten (HOG's). De HOG's kunnen meerdere mentale staten tegelijk representeren, waardoor ze als één enkele bewustzijnservaring ervaren worden. Daarnaast refereren HOG's niet alleen maar naar de mentale staten die ze representeren, maar ook naar zichzelf als diegene die de mentale staat heeft. Hierdoor lijkt het alsof de mentale staten worden waargenomen door een temporeel continue zelf. De tweede claim die ik maak is dat het gevoel van psychologische vrijheid geen objectief feit is. Zonder mij te mengen in het traditionele ontologische debat over vrije wil, zal ik beargumenteren dat het psychologische gevoel van vrijheid ook verklaard kan worden door de HOG theorie, zonder een temporeel continue zelf te hoeven postuleren. Ik concludeer dus dat het formele zelf van Searle een onnodig gepostuleerde entiteit is.

## Inhoudsopgave

<b>Inleiding</b> .....	<b>1</b>
<b>Persoonlijke identiteit en het zelf</b> .....	<b>3</b>
De ego-theorie van persoonlijke identiteit .....	4
De bundeltheorie van persoonlijke identiteit .....	5
Bundeltheorieën van persoonlijke identiteit in de westerse analytische filosofie .....	6
<b>John Searle</b> .....	<b>8</b>
Searle's problemen met persoonlijke identiteit .....	8
De vier criteria van persoonlijke identiteit .....	10
Het bewustzijnsveld .....	13
Het zelf als formele entiteit .....	15
Searle's valkuilen .....	20
<b>De HOG theorie</b> .....	<b>21</b>
De HOG theorie en de neurowetenschap .....	23
Objecties tegen de HOG theorie .....	24
<b>Kritiek op het zelf als de binder van ervaringen</b> .....	<b>26</b>
De HOG theorie als verklaring voor de gevoelde eenheid van ervaring .....	27
De HOG theorie als alternatief voor een temporeel continue formele zelf .....	29
<b>Kritiek op het zelf als de maker van keuzes</b> .....	<b>32</b>
Searle's illusionaire gat van psychologische vrijheid .....	33
Het gevoel van psychologische vrijheid verklaard .....	37
Initiëren van acties, reflectievermogen en het repertoire van herinneringen .....	40
<b>Conclusie</b> .....	<b>41</b>
<b>Referenties</b> .....	<b>45</b>

Nurture your mind with great thoughts, for you will never go higher  
than you think.

*Benjamin Disraeli*



## Inleiding

Het plot van de serie 'Doctor Who?' geeft ons een scenario waarin verschillende personages met elk hun eigen voorkeuren en herinneringen toch allemaal een en dezelfde incarnatie zijn van de 'Doctor'. Hoe is dit mogelijk? Hoe kunnen verschillende personages allemaal dezelfde 'Doctor' zijn als ze zoveel van elkaar verschillen in zoveel verschillende opzichten?

Dit is misschien een irreëel voorbeeld uit een sci-fi serie, maar je kunt dezelfde vragen ook over jezelf stellen. Als je erbij stilstaat is door de jaren heen nagenoeg niks wat jou kenmerkt hetzelfde gebleven. Bijna elke cel in jouw lichaam is wel een keer vervangen door een nieuwe, de vorm van jouw lichaam is veranderd, jouw stem is vervormd, je hebt andere voorkeuren en overtuigingen gekregen, je hebt nieuwe herinnering opgedaan, je bent oude herinneringen vergeten en ga zo maar door. Je kunt je dus afvragen wat jou nu precies jou maakt. Ondanks deze onbeantwoorde vraag gebruiken we in het dagelijks leven het woord 'ik' zonder er echt bij stil te staan. Zinnen zoals 'ik voel me goed', 'ik heb zin om te lezen', en 'ik zie een mooie vogel' worden gedachteloos gebruikt tijdens conversaties. Deze 'ik' lijkt vanzelfsprekend te zijn, maar als we op zoek gaan naar deze 'ik' en dat waar het naar verwijst blijkt het toch moeilijker te duiden dan in eerste instantie wordt gedacht.

John Searle is een analytisch filosoof die zich bezighoudt met vraagstukken omtrent persoonlijke identiteit en het zelf. Hij beweert dat de meeste filosofen het er over eens zijn dat Hume het laatste woord over persoonlijke identiteit en het zelf heeft gezegd (Searle 2004, 279). Hume beargumenteerde dat het zelf niet bestaat, omdat onze ervaring van de werkelijkheid niks meer is dan een opeenvolging van percepties. Als we in deze percepties op zoek gaan naar het zelf, vinden we niks (Hume 1978, 252). Searle is echter één van de filosofen die het niet eens met deze conclusie. Hij stelt dat er naast de sequentie van percepties wel een zelf gepostuleerd moet worden om onze ervaring van de werkelijkheid te

kunnen verklaren. Wij kunnen namelijk, volgens Searle, zonder een zelf te postuleren niet verklaren waarom (1) de werkelijkheid wordt ervaren als één enkelvoudige bewustzijnsstaat en waarom (2) wij het gevoel hebben dat we op elk moment van de tijd vrije keuzes kunnen maken. Searle stelt dus dat de conclusie van Hume, dat het zelf niet bestaat uit percepties, correct is, maar dat de conclusie dat deze zelf helemaal niet bestaat, niet klopt. Het zelf bestaat volgens Searle namelijk wel in een formele vorm. Het formele zelf bestaat niet uit percepties, maar het heeft wel de causale macht om de percepties te beïnvloeden. Deze formele zelf heeft een aantal eigenschappen. Het neemt bewust de werkelijkheid waar, maakt vrije keuzes en heeft de mogelijkheid om op deze keuzes te reflecteren.

Er is veel kritiek geleverd op Searle's theorie van persoonlijke identiteit en het zelf. Dennett stelde bijvoorbeeld dat het zelf, net als bewustzijn, niks meer is dan een hersenspindel gecreëerd door het brein. Het zelf van Searle, ook al is het een formele notie, is niks meer dan een fata morgana zonder enige ontologische waarde (Dennett 2003, 39-40). Kaufmann beargumenteerde in het verlengde hiervan dat het formele zelf van Searle solipsistisch van aard is en daarom moeilijk te verdedigen (Kaufmann 2005, 456). Wat de verschillende kritieken op Searle's formele notie van het zelf echter niet goed kunnen verklaren, is waarom het voor ons fenomenologisch toch voelt alsof we een temporeel continue zelf zijn die in het middelpunt van onze ervaring van de werkelijkheid staat. De vragen die ervoor zorgden dat Searle tot de conclusie kwam dat er een formele zelf bestaat, blijven onbeantwoord.

In deze thesis probeer ik op deze vragen van Searle een antwoord te geven. Ik zal beargumenteren waarom het niet nodig is om een zelf te postuleren om te begrijpen waarom wij de werkelijkheid als één bewustzijnservaring ervaren en waarom het voor ons voelt alsof we vrije keuzes maken. Hierin zal ik vooral gebruik maken van de hogere orde gedachten (HOG) theorie, in de traditie van Rosenthal. De HOG theorie stelt dat bewuste staten mentale staten zijn die op een bepaalde



manier in relatie staan met een mentale staat van een hogere orde (Rosenthal 2002, 235-6). Met deze HOG theorie kunnen de eigenschappen die Searle toeschrijft aan het formele zelf ook verklaard worden. De HOG theorie wordt ondersteund door neurowetenschappelijke onderzoeken en is naar mijns inziens een theorie die in lijn is met hoe wij de werkelijkheid ervaren. Waarom dit zo is, zal ik in deze thesis expliciteren.

Ik zal beginnen met een korte introductie te geven van het debat omtrent het thema persoonlijke identiteit en het zelf. Het is een debat dat een lange geschiedenis heeft, die we moeten uitwerken om te begrijpen waar Searle in het debat staat. Daarna zal ik Searle's argumentatie systematisch uitwerken en zal ik de begrippen die Searle gebruikt verder uitleggen. Nadat ik Searle's positie uiteen heb gezet, zal ik op de twee hoofdargumenten van Searle, voor het bestaan van een formele zelf, kritiek leveren. Dit zal ik, zoals eerdergenoemd, vooral doen door te verwijzen naar de HOG theorie. Ik zal daarom eerst een samenvatting geven van de HOG theorie en vervolgens ook beargumenteren waarom ik denk dat de HOG theorie een standvastige theorie is die ondersteund wordt door neurowetenschappelijke onderzoeken.

## **Persoonlijke identiteit en het zelf**

Vragen zoals 'wat ben ik?', 'wie ben ik?' en 'wanneer begon ik te bestaan?' blijken, zoals eerder benoemd, lastig te beantwoorden vragen. Hoe kan het dat we moeite hebben met het uitleggen wat de 'ik' is als we het in het dagelijks zo vaak gebruiken? Waarom is deze 'ik' zo lastig te vangen? Vraagstukken zoals deze vallen onder het thema persoonlijke identiteit en het zelf. De centrale vraag waar dit thema zich mee bezig houdt is: waarom lijkt het alsof ik een enkele, spatio-temporeel continue zelf ben die in het middelpunt van de ervaringen staat? (Blackmore 2013, 106).

Over de jaren heen hebben vele filosofen en wetenschappers geprobeerd een logisch sluitend antwoord te vinden op deze hoofdvraag. De mogelijke theorieën die hierover gepostuleerd werden lopen ver uiteen. Parfit (1987, 2-3) categoriseerde deze en deelde ze op in twee verschillende groepen, de ego-theorieën en de bundeltheorieën van persoonlijke identiteit.

### **De ego-theorie van persoonlijke identiteit**

Volgens de ego-theorie kan de spatio-temporele continuïteit van een persoon alleen maar verklaard worden als je het bestaan van een ego of subject postuleert dat aan het begin staat van de ervaringen. Een ego-theoreticus claimt dat alle individuele ervaringen en indrukken uit de buitenwereld op elk moment van de tijd worden ervaren door een continue zelf, subject of persoon. Deze zelf zorgt ervoor dat alle informatie die binnenkomt via de verschillende zintuigen aan elkaar gekoppeld wordt. Hierdoor nemen we de wereld als een eenheid waar (Parfit 1987, 2).

René Descartes is een van deze ego-theoretici. Hij gebruikte de methode van radicale twijfel om erachter te komen wat onbetwifelbaar en daardoor noodzakelijk waar was. Aan alle ervaringen kon hij twijfelen, behalve aan het feit dat er een zelf of bewustzijn aan het twijfelen was. Hij trok daarom de conclusie dat het zelf aan de basis van de ervaring staat (Descartes 2018, 60). 'Ik denk, dus ik ben', luidt het citaat. Hij claimde dat dit denkende zelf, de *res cogitans*, op zichzelf kan bestaan en dat het niets anders nodig heeft dan zichzelf om dit te doen. Deze *res cogitans* stond in contrast met de *res extensa*, waarmee hij de ruimtelijke externe wereld bedoelde. Beide bestonden volgens Descartes uit een verschillende substantie (Descartes 1972, 102). Deze Cartesiaanse visie van het zelf is een intuïtieve theorie die in lijn staat met hoe wij onszelf voelen. Uit psychologische onderzoek blijkt dat kinderen vanaf drie jaar oud al het bestaan van een ziel postuleren (Bloom 2004, 201).

John Locke is een mede ego-theoreticus die voortbouwde op het werk van Descartes. Locke praatte over de 'persoon' waar Descartes over de 'zelf' sprak. Deze twee termen zijn in dit geval echter aan elkaar gelijk te stellen, omdat ze niet afwijken in betekenis (Nimbalkar 2011, 1-2). Locke beschreef een persoon als een 'denkend intelligent wezen, dat kan redeneren en reflecteren, en die bewust kan worden van zichzelf, hetzelfde denkende ding, op verschillende tijden en plaatsen' (Locke 1975, 335). Waar Locke afwijkt van Descartes is dat Locke meer de nadruk legt op het temporele aspect van de persoon. Volgens Locke is een persoon namelijk alleen maar een persoon tot in hoeverre hij zichzelf bewust kan herinneren. Persoonlijke identiteit is daarom, volgens Locke, onlosmakelijk verbonden met het geheugen.

De ego-theorie, zoals ik die hierboven heb beschreven, is een intuïtieve theorie die lijkt aan te sluiten bij hoe wij de werkelijkheid ervaren. Er zijn echter vele filosofen en wetenschappers die kritiek hebben op deze intuïtieve notie van het zelf. Deze stellen dat dit ego, of het subject van de ervaring, niet nodig is om te kunnen verklaren waarom we de werkelijkheid waarnemen zoals we doen. Tegenstanders van de ego-theorie postuleerden een alternatieve theorie van het zelf, genaamd de bundeltheorie.

### **De bundeltheorie van persoonlijke identiteit**

Volgens de bundeltheorie is het niet nodig om een ego te veronderstellen om te begrijpen waarom we onze werkelijkheid ervaren als een eenheid. David Hume, een van de grondleggers van deze theorie, ontkende het bestaan van een permanente en immateriële zelf. Hij stelde dat de mens niks meer is dan een bundel van percepties die onvoorstelbaar snel achter elkaar plaatsvinden (Hume 1978, 252). Al onze sensaties, impressies en ideeën lijken aan elkaar gebonden te zijn door de relaties tussen de ervaringen en de herinneringen die we er later van maken. Buiten deze bundel van percepties om hoeft er geen ego gepostuleerd te

worden om te begrijpen hoe we de werkelijkheid ervaren. Er is volgens Hume dus geen zelf die alle ervaringen samenvoegt.

Hume kwam tot deze conclusie doormiddel van introspectie. Hij probeerde in zichzelf op zoek te gaan naar zichzelf, maar kwam tot de conclusie dat, wanneer hij dit probeerde, hij niks meer tegenkwam dan percepties. Hij voelde zijn tenen tegen de binnenkant van zijn schoenen schuren, voelde een lichtelijke hoofdpijn en had het een beetje koud, maar buiten deze percepties om vond hij niks dat kon duiden op een immateriële permanente zelf. Hij concludeerde daarom dat het zelf een bundeling van deze ervaringen is en niet iets dat daarbuiten staat. Wanneer deze percepties er niet zijn, zoals tijdens de diepe slaap, dan is het zelf er ook niet (Hume 1978, 252).

Je zou deze notie van het zelf, zoals Hume het bedoelt, kunnen vergelijken met een samenleving die opgebouwd is uit verschillende mensen. Het zelf is, net als een samenleving, niet een permanent ding dat op zichzelf staat. Het is een collectieve naam voor alle percepties samen, net zoals de samenleving een collectieve naam is voor alle deelnemers van een maatschappij. Deze percepties zijn met elkaar verbonden volgens de wetten van de natuur, net zoals mensen in een samenleving verbonden zijn door rechten en plichten.

### **Bundeltheorieën van persoonlijke identiteit in de westerse analytische filosofie**

Deze bundeltheoretische opvatting van het zelf van Hume werd door veel westerse analytische filosofen overgenomen. William James was hier een van. Volgens James klopt Hume's conclusie dat er geen 'spiritueel wezen' of een ego bestaat (James 1890, 370). Hume gaat echter met zijn theorie van het zelf tegen het gezond verstand van de gehele mensheid in. Hij kan volgens James niet goed verklaren waarom het dan toch voor ons voelt alsof we een samenhangend continu subject zijn dat aan de basis ligt van al onze ervaringen. Iemand die bewustzijn als eerste principe uit de vergelijking haalt, moet alsnog de functies die het bewustzijn lijkt te hebben kunnen verklaren (James 1904, 478).

James formuleert hier een antwoord op. Volgens James (1890, 401) staan onze gedachtes gelijk aan de denker. Er is geen speciale denker die achter de werkelijkheid van percepties staat die de gedachtes denkt. De denker is niks meer dan een gedachte van een speciaal soort, die hij 'de Gedachte' noemt. Op elk moment van de tijd is er volgens James een speciale gedachte die de voorgaande gedachtes herinnert en claimt dat die voorgaande gedachtes een onderdeel zijn van zichzelf. Deze Gedachte beoordeelt bij opkomst welke gedachtes bij hem horen en welke niet. Na een bepaalde tijd verdwijnt de Gedachte en komt er een nieuwe Gedachte op die hetzelfde proces doorloopt als de voorgaande Gedachte (James 1890, 363).

Bertrand Russell deelde de opvatting van James. Hij stelde dat er buiten de bundel van percepties geen metafysische zelf of ego bestaat. De mentale continuïteit van een persoon kan volledig worden verklaard door gewoontes en het geheugen alleen. Hij beargumenteerde dat het woord 'ik' niks meer is dan een grammaticale handigheid (Russell 1966, 550). De 'ik denk' in het citaat 'ik denk, dus ik ben' van Descartes zou volgens Russell vervangen moeten worden met 'er zijn gedachten'. Dit zou de onwaarheid van Descartes' conclusie aan het licht doen brengen.

Ludwig Wittgenstein sprak zich in de *Tractatus* ook duidelijk uit over de onwaarheid van het bestaan van een subject achter de percepties van de wereld. Er is volgens Wittgenstein geen subject dat denkt of ideeën toegespeeld krijgt. Het subject is niks meer dan een oppervlakkige psychologische stelling die geen enkele onderbouwing ondervindt (Wittgenstein 1992, 5.631).

Alfred Ayer sloot zich ook aan bij James, Russell en Wittgenstein. Hij zei dat er geen spiritueel denkende substantie bestaat buiten de bundel van percepties om. Zelfbewustzijn staat volgens Ayer gelijk aan de claim dat een bepaalde ervaring tot jou toe behoort (Ayer 1973, 120). Dit sluit aan bij de thesis van James die stelde dat er een speciale discriminerende Gedachte bestaat die bepaald wat bij zichzelf hoort en wat niet. Ook denkt Ayer, net als Russell en Wittgenstein, dat

de meeste filosofische problemen voortkomen uit de complexiteit van de taal. Je kunt niet van taal uitgaan om bij feiten te komen. Taal is namelijk geen veilig instrument om tot de waarheid te komen. Zo hebben 'ik' en 'mijn lichaam' verschillende grammaticale betekenissen in de taal. 'Mijn lichaam' geeft een bezit aan, terwijl 'ik' een eigennaam is. Dit onderscheid in de taal is volgens Ayer echter onrechtvaardig en kan daarom ook niet worden aangewezen als bewijs voor het bestaan van een metafysische entiteit buiten het lichaam om (Ayer 1973, 120-21).

## **John Searle**

John Searle is een Amerikaanse analytisch filosoof die actief bijdraagt aan de literatuur in de taal filosofie, cognitiefilosofie en sociale filosofie. Hij is voornamelijk bekend geworden door zijn werk over 'speech acts' waarin hij beïnvloed werd door het werk van Austin en Wittgenstein. Daarnaast schrijft hij ook uitgebreid over de onderwerpen mind, intentionaliteit, bewustzijn, vrije wil, rationaliteit, persoonlijke identiteit en het zelf. In deze thesis richt ik mij echter vooral op zijn werk over persoonlijke identiteit en het zelf, al komen de rest van de onderwerpen waarover hij schrijft hierbij ook aan bod.

### **Searle's problemen met persoonlijke identiteit**

Volgens Searle zijn er drie problemen van persoonlijke identiteit die opspelen bij het omschrijven van het zelf.

1. Aan welke criteria moet persoonlijke identiteit voldoen?
2. Moeten we een extra zelf postuleren, buiten het lichaam en de sequentie van percepties om, om onze ervaring van de werkelijkheid te kunnen verklaren?
3. Wat geeft mij mijn identiteit en persoonlijkheid?

Vraag 3 gaat volgens Searle over hoe de identiteit van iemand wordt gevormd door de cultuur en maatschappij. Identiteit wordt hierin vooral gezien als een combinatie van karaktereigenschappen en persoonlijkheid. Deze identiteit staat daarom volgens Searle los van het metafysische probleem van het bestaan van een zelf (Searle 2005, 280-82). Searle houdt zich in zijn werk vooral bezig met vraag 1 en 2 en stelt dat vraag 3 een vraag is voor psychologen en sociologen om te beantwoorden.

Alhoewel het probleem omtrent identiteit al zo oud is als de filosofie zelf, blijft het een puzzel, volgens Searle. Zo beschrijft hij het dilemma van het 'schip van Theseus' om zijn punt duidelijk te maken. Dit dilemma gaat als volgt: stel je vaart met een schip die volledig van houten planken is gemaakt op de Middellandse Zee. Plank voor plank wordt het schip tijdens het varen vernieuwd. Aan het einde van de reis is elke plank vervangen door een nieuwe. Is dit dan nog steeds hetzelfde schip van Theseus? De meeste mensen zullen hier intuïtief ja op zeggen. De identiteit van een schip wordt namelijk bepaald door zijn functie. Als het schip blijft varen is het hetzelfde schip. Maar stel nu dat er iemand was die alle planken van het schip van Theseus verzamelde en hier een volledig nieuw schip van maakte. Welk van de twee schepen is het echte schip van Theseus? Welk kenmerk is het belangrijkste om de identiteit van het schip te bepalen? De continuïteit van de functie of de continuïteit van de onderdelen waar het uit bestaat? Volgens Searle is hier geen eenduidig antwoord op te geven. Er is zijn geen feitelijke vragen die kunnen bepalen welk schip het echte schip is. Het is aan ons om te bepalen welke schip het echte schip van Theseus is (Searle 2016).

Niet alle problemen omtrent persoonlijke identiteit zijn echter zo makkelijk op te lossen als het traditionele probleem van het schip van Theseus. Volgens Searle voelen we alsof er een 'speciaal' probleem blijft bestaan, zelfs als we een logisch antwoord vinden op vragen zoals die van het schip van Theseus. We voelen onszelf namelijk altijd onszelf vanuit een eerste-persoonsperspectief en dit is

essentieel voor onze identiteit, stelt Searle. We kunnen ons bijvoorbeeld allemaal inbeelden hoe het is om wakker te worden in een ander lichaam. Stel dat er een breintransplantatie plaatsvindt, waarin je brein in het lichaam van iemand anders wordt geplaatst. Dan nog zul je het gevoel hebben dat je vanuit eerste-persoonsperspectief dezelfde persoon bent als diegene die je was voor de breintransplantatie, volgens Searle.

### **De vier criteria van persoonlijke identiteit**

Om de identiteit van iemand vast te stellen maken we volgens Searle gebruik van vier criteria. Het eerste criterium, en tevens het belangrijkste, is het lichaams criterium. Iemand is dezelfde persoon in het hier en nu als diegene die hij was bij de geboorte, omdat het lichaam door de tijd heen continu dezelfde ruimte in heeft genomen. Dit kan echter op zichzelf niet persoonlijke identiteit verklaren, zegt Searle. Door de jaren heen veranderen nagenoeg alle micro-onderdelen van het lichaam, zoals de moleculen en cellen waaruit het is opgebouwd. Er speelt dus bij het bepalen van persoonlijke identiteit meer mee dan alleen het lichaam.

Het tweede criterium van persoonlijke identiteit is de relatieve temporele continuïteit van het lichaam. Zelfs na veranderingen aan het lichaam door de tijd heen, blijf je spreken van een en dezelfde persoon. Het hierboven genoemde probleem kun je oplossen door te stellen dat persoonlijke identiteit afhangt van hoe snel de veranderingen plaatsvinden. Er is een bepaald natuurlijk ritme waarop een lichaam vanaf de geboorte verandert en groeit. Als hiervan plotseling wordt afgeweken, doordat het bijvoorbeeld in één keer verandert in het lichaam van een insect, zou je kunnen spreken van een andere identiteit. Dit criterium, samen met het lichaams criterium, zijn criteria ten opzichte van het derde-persoonsperspectief. Dit is hoe vanuit derde-persoon de persoonlijke identiteit van een persoon wordt vastgesteld. Deze kunnen echter niet het eerste-persoonsperspectief van persoonlijke identiteit verklaren. Als je in één keer verandert in een insect, dan zullen misschien andere mensen je misschien niet



meer dezelfde persoonlijke identiteit toeschrijven, maar toch blijf je altijd vanuit eerste-persoonsperspectief je jezelf voelen. Om dit eerste-persoonsperspectief van persoonlijke identiteit te kunnen verklaren heb je een geheugen nodig volgens Searle.

Het derde criterium van persoonlijke identiteit is dan ook het geheugencriterium. Locke stelde dat het meest essentiële aan persoonlijke identiteit, het bewustzijn is (Locke 1975, 330). Naast het lichaam is ook bewustzijn nodig om persoonlijke identiteit te kunnen verklaren. Het zou namelijk kunnen zijn dat je plotseling wakker wordt in een ander lichaam en je toch volledig jezelf voelt. De sequentie van ervaringen vinden in het nieuwe lichaam nog steeds opeenvolgend achter elkaar plaats en de ervaringen in het hier en nu zijn nog steeds gekoppeld aan de herinneringen van ervaringen uit het verleden. Volgens Searle bedoeld Locke met bewustzijn onze ervaringen van het geheugen in het hier en nu, die gericht zijn op de continuïteit tussen de huidige zelf en het zelf uit het verleden, die ervaringen had waarop het geheugen in het hier en nu is gebaseerd (Searle 2005, 9). Met andere woorden, het bewustzijn van Locke staat gelijk aan wat we tegenwoordig het lange termijn geheugen noemen.

Searle stelt net als Locke dat het gevoel dat we door de tijd heen continu dezelfde persoon blijven grotendeels te danken is aan het geheugen. Hij neemt het voorbeeld van Leibniz over om zijn punt duidelijk te maken. Stel je voor dat je de nieuwe keizer van China wordt, maar dat je het hele leven dat je tot nu toe hebt geleefd vergeet. Dat je geen enkele herinnering meer in je geheugen hebt opgeslagen van het moment voordat je gekroond werd. In dit voorbeeld bij jij als persoon opgehouden te bestaan en is een nieuwe keizer van China geboren. De continuïteit van het bestaan van een persoon is volgens Searle dus afhankelijk van het geheugen.

Searle stelt echter dat het geheugen alleen ook niet persoonlijke identiteit in zijn volledigheid kan verklaren. De stelling dat dit wel kan is namelijk circulair van aard. We kunnen alleen zeggen dat de herinneringen van een persoon

daadwerkelijk rechtmatig aan diegene toebehoren als de persoon identiek is aan diegene die de herinneringen in het verleden heeft opgedaan (Searle 2004, 288-291). De correctheid van de herinneringen veronderstelt dus dat de persoon door de tijd heen dezelfde is gebleven. Het is daarom niet mogelijk om het geheugen te gebruiken om de persoonlijke identiteit vast te leggen.

Een vierde criterium van persoonlijke identiteit is de continuïteit van persoonlijkheid. Deze is volgens Searle minder belangrijk dan de voorgaande drie, maar toch zou je kunnen zeggen dat, net als het lichaam, je persoonlijkheid relatief stabiel blijft. Door de jaren heen veranderen wellicht je voorkeuren en disposities, maar ook deze veranderingen vinden gradueel plaats (Searle 2004, 287-88). Als je morgen wakker zou worden en je zou jezelf gedragen en voelen alsof je Michael Jackson bent dan zou men zich kunnen afvragen of je nog steeds dezelfde persoon bent als daarvoor.

Deze vier criteria van persoonlijke identiteit geven ons inzicht in wat het maakt dat een persoon dezelfde persoon blijft door tijd en veranderingen heen. Deze criteria werken in het alledaagse leven goed genoeg om antwoord te kunnen geven op de vraag wat wij normaal gesproken beschouwen als persoonlijke identiteit. Het eerste probleem van persoonlijke identiteit is hiermee volgens Searle opgelost. De vraag waar echter nog een antwoord op rest te geven is de vraag waar deze vier criteria nu precies naar verwijzen. Staat het gevoel van persoonlijke identiteit gelijk aan de sequentie van percepties, zoals Hume die bedoelde, of moet er toch iets extra's worden gepostuleerd om onze ervaring van het zelf te kunnen verklaren? Searle concludeert van wel. Het zelf dat hij postuleert is echter geen zelf die te vinden is in percepties. Het zelf is volgens Searle een puur formele entiteit (Searle 2016).

Voor ik Searle's theorie van het zelf verder uiteen kan zetten moet eerst beschreven worden wat Searle precies met bewustzijn en het bewustzijnsveld bedoeld. Searle ziet het zelf namelijk als waarnemer, manipulator en organisator van dit bewustzijnsveld (Searle 2005, 16-17).

## **Het bewustzijnsveld**

Het probleem omtrent de notie van bewustzijn is een al lang bestaand probleem dat tegenwoordig te reduceren is tot de vraag: hoe kunnen onbewuste materialistische deeltjes in het brein zorgen voor bewustzijnservaringen? Op deze vraag zijn door filosofen diverse metafysische antwoorden gegeven. Searle (2000, 559) ziet bewustzijn echter niet als een metafysisch, maar als een biologisch fenomeen, die in hetzelfde rijtje past als de werking van het verteringskanaal, de ademhaling en fotosynthese. Het brein creëert een bewustzijnsveld zoals de maag en de darmen vertering veroorzaken (Searle 2005, 12). Bewustzijn is volgens Searle dus een neurobiologisch proces. Hij bedoelt hiermee niet dat mentale staten gelijk staan aan breinstaten, zoals de identiteitstheorie het wel beschrijft. Searle stelt namelijk dat we neurobiologische processen zo kunnen herdefiniëren dat het ook bewustzijn omvat (Searle 2005, 125-6). Op deze manier probeert Searle een omweg te vinden voor onder andere het lichaam-geest probleem. Deze herdefinitie van neurobiologische processen zegt dus niet dat eerste-persoonservaringen van bewustzijn gelijk staan aan derde-persoons neurobiologische processen, maar dat ze er wel een onderdeel van zijn. Searle benoemt deze manier van kijken naar mentale staten 'biologisch naturalisme'.

Bewustzijn kan volgens Searle worden gedefinieerd als 'innerlijke staten van gevoelens, waarneming en alertheid dat meestal begint bij het opstaan uit een droomloze slaap, door de dag heen in meer of mindere mate aanwezig blijft, en stopt wanneer men weer gaat slapen, in coma belandt, doodgaat of op een andere manier 'onbewust' wordt' (Searle 2005, 10). Searle (2000, 560) beschrijft dat dit bewustzijn drie eigenschappen bezit waar elke theorie van bewustzijn een verklaring voor zou moeten kunnen geven: het is kwalitatief, subjectief en wordt gekenmerkt door eenheid.

Het is kwalitatief, omdat elke bewustzijnsstaat kwalitatief gevoeld wordt. De ervaring van het voelen van regen op de huid is anders dan die van het horen van klassieke muziek. Beide ervaringen hebben een ander kwalitatief karakter.

Searle neemt het begrip van kwalitatieve karakter over van Nagel. Volgens Nagel kun je bij elk bewust wezen afvragen hoe het is om dat wezen te zijn (Nagel 1974, 436). Bij een niet-bewust object, zoals een fiets of een stoeptegels, kun je dus niet afvragen hoe het is om dat object te zijn. Er zijn filosofen die het kwalitatief karakter van bewustzijn aanduiden met het woord '*qualia*' en claimen dat dit een speciaal probleem van bewustzijn vormt (Chalmers 1996, 247-49). Searle claimt echter dat qualia niks anders is dan een verzamelnaam voor bewustzijnsstaten (Searle 2000, 561). Volgens Searle bestaan gedachtes ook uit qualia, ook al zijn er filosofen die claimen van niet.

Het is subjectief, omdat bewustzijnsstaten alleen kunnen bestaan als ze ervaren worden door een menselijk of dierlijk subject (Searle 2000, 561). Volgens Searle impliceert het kwalitatief karakter van bewustzijn het subjectief karakter ervan, omdat iets alleen kwalitatief gevoelt kan worden als er een subject is dat de qualia ervaart. Subjectieve bewustzijnsstaten hebben dus altijd een eerste-persoonsontologie, omdat ze altijd worden ervaren door een subject. Ze worden ervaren door een 'ik' die de ervaringen heeft. Searle claimt dat dit ook de reden is waarom bewustzijn niet objectief onderzocht kan worden. Bewustzijn is altijd gekoppeld aan een subject, je praat over *mijn* bewustzijn of *jouw* bewustzijn, maar nooit over bewustzijn als een objectief feit (Searle 1992, 93-4).

Bewustzijn wordt daarnaast gekenmerkt door eenheid. Alle bewuste ervaringen die een subject heeft in zijn leven zijn allemaal onderdeel van één enkelvoudig bewustzijnsveld (Searle 2000, 561). Alle verschillende bewustzijnsstaten die je op een bepaald moment in de tijd hebt worden allemaal ervaren als één samenhangende bewuste ervaring. De afzonderlijke ervaringen van bijvoorbeeld de druk die je lichaam uitoefent op de stoel, de geur van de vers gezette koffie en het geluid van de wind die buiten door de bomen raast, vinden allemaal plaats in één bewustzijnsveld. Searle stelt dat net zoals het kwalitatief karakter van bewustzijn het subjectief karakter impliceert, dat de eenheid van het bewustzijnsveld het kwalitatief en subjectief karakter impliceert. Deze drie

eigenschappen van bewustzijn zijn dus allemaal met elkaar verbonden en de ene impliceert de andere.

Dit bewustzijnsveld kan op een aantal manier gemanipuleerd worden, volgens Searle. Als eerste kunnen we door onze wil de aandacht in het bewustzijnsveld verschuiven. Zonder je hoofd of je ogen te bewegen kun je alsnog je aandacht richten op verschillende delen van het visuele veld. Ten tweede kunnen we door deze zelfde wil ook het hele bewustzijnsveld in een keer veranderen door bijvoorbeeld op te staan of ons hoofd te bewegen.

Het probleem dat hier echter uit volgt is, wie of wat verschuift deze aandacht? Wie verandert het bewustzijnsveld door op te staan? En wie ervaart dit bewustzijnsveld? Deze vragen kunnen volgens Searle niet beantwoord worden als we geen zelf postuleren. Searle claimt daarom dat er meer is dan alleen een sequentie van percepties die samen het bewustzijnsveld vormen. Er is een zelf die dit bewustzijnsveld kan manipuleren, organiseren en waarnemen. Deze zelf is echter geen zelf zoals Hume die omschrijft, maar een puur formele entiteit. Wat dit precies inhoudt zal ik in de volgende paragraaf uitleggen.

### **Het zelf als een formele entiteit**

Searle claimt dat de meeste filosofen het er over eens zijn dat Hume het laatste woord over persoonlijke identiteit en het zelf heeft gezegd (Searle 2004, 279). Ze concluderen dat er geen zelf gepostuleerd hoeft te worden om te begrijpen hoe wij de werkelijkheid ervaren, omdat er niks meer bestaat dan een opeenvolging van percepties. Hume zocht naar een zelf die alle percepties bij elkaar houdt, maar hij vond er geen een. Searle claimt dat Hume hier een logische conclusie trekt en geen psychologische conclusie. Logisch gezien kan het zelf niet worden teruggevonden in de percepties, omdat elke perceptie of ervaring die je hebt, ook al zijn dit ervaringen die een leven lang duren, niks meer zijn dan andere percepties (Searle 2004, 291-92). Stel je voor dat je jouw hele leven lang een bepaalde vlek in je visuele veld ziet, is dit dan het zelf? Nee, concludeert Searle, er

is geen enkele perceptie of ervaring die voldoet aan de eisen om een zelf genoemd te kunnen worden. Het argument van Hume is overtuigend, maar laat wel iets fundamenteels weg (Searle 2016). We kunnen met alleen de sequentie van percepties namelijk niet verklaren waarom wij de werkelijkheid ervaren als een eenheid. Daarnaast hebben we volgens Searle ook altijd een gevoel van psychologische vrijheid dat door deze redenering niet kan worden weg verklaard. Om onze ervaring van de realiteit in zijn totaliteit te kunnen verklaren, moeten we volgens Searle dus wel een zelf postuleren. Deze zelf is echter geen metafysische zelf, maar een puur formele entiteit die niet bestaat uit percepties.

Het eerste argument dat Searle aanhaalt voor het bestaan van een formele zelf gaat over de ervaring van ons bewustzijnsveld. Zoals eerder beschreven staan de ervaringen die we hebben van de werkelijkheid niet los van elkaar. Ze worden allemaal ervaren als één enkelvoudige bewustzijnservaring. Zo lees je op dit moment deze thesis, hoor je wellicht buiten de vogels fluiten en ruik je de net vers gemaakte appeltaart. Deze bewustzijnservaringen vormen allemaal jouw bewustzijnsveld in het hier en nu. Searle (2004, 292) claimt dat deze ervaringen allemaal met elkaar verbonden worden door het zelf. Zonder een zelf die deze ervaringen bindt, zouden we namelijk de wereld ervaren als afzonderlijke bewustzijnsstaten. Searle postuleert daarom een formele entiteit die hij ziet als de binder van ervaringen.

Hij stelt daarnaast dat de continuïteit van het bewustzijnsveld door de tijd heen wordt geclaimd door deze formele zelf, die het bewustzijnsveld ziet als de continuïteit van zijn of haar eigen bewustzijn (Searle 2004, 293). Hiermee bedoelt hij dat elke bewustzijnsstaat door de tijd heen wordt ervaren als een onderdeel van één continue ervaring. De bewustzijnsstaat die je nu ervaart wordt niet ervaren als losstaand van de bewustzijnsstaat die je een half uur, een maand of zelfs een jaar geleden ervoer. Het formele zelf, die voor de eenheid van ervaringen zorgt, wordt dus ook gekenmerkt door temporele continuïteit. Het formele zelf

blijft dus door de tijd heen hetzelfde formele zelf, indien het aan Searle's criteria van persoonlijke identiteit voldoet.

Een tweede argument voor het bestaan van een formele zelf gaat over het hebben van vrije keuze, het maken van keuzes en de redenen waarom we bepaalde keuzes maken. Het feit dat het mogelijk is om de aandacht te verschuiven binnen het bewustzijnsveld of zelfs het hele bewustzijnsveld te veranderen door het toepassen van de wil, is voor Searle een andere belangrijke aanwijzing voor het bestaan van een zelf. Er zou ook een realiteit mogelijk kunnen zijn waar geen actieve zelf bestaat die kan handelen uit vrije keuze. Deze passieve zelf zou dan alleen maar passieve waarneming ondervinden en zou geen causale macht hebben op dat wat er gebeurt (Searle 2005, 15).

Wanneer je bewust vrijwillige acties onderneemt, zoals het verschuiven van je aandacht in het bewustzijnsveld, heb je altijd een gevoel van psychologische vrijheid. Je hebt dan het gevoel dat je uit eigen vrije keuze een handeling hebt verricht, terwijl het ook mogelijk had kunnen zijn dat je in dezelfde situatie iets anders had gedaan.

'I have the sense I am doing this, but I could, right here and right now, be doing something different. In such cases I have the impression that the causes of my action, in the form of the reasons on which I am acting, are not causally sufficient to determine the action' (Searle 2005, 15).

In normale situaties kun je dus volgens Searle alleen maar praten over gemotiveerde handelingen, niet over gedetermineerde handelingen. Motivaties, overtuigingen en verlangens functioneren namelijk alleen als mogelijke redenen voor het uitvoeren van handelingen, niet als de causale oorzaak ervan. Zo maakt Searle een onderscheidt tussen stellingen als:

1. Ik kreeg buikpijn, omdat ik voor Bush wilde stemmen.
2. Ik tekende een X op het stemformulier, omdat ik voor Bush wilde stemmen.

Vanuit de standaard interpretatie van deze zinnen zou je kunnen zeggen dat stelling 1 causaal bepalende condities heeft. Je kreeg de buikpijn, omdat je voor Bush wilde stemmen. Stelling 2 is echter een ander verhaal. Volgens Searle tekende je een X op het stemformulier, maar had je in dezelfde situatie ook andere dingen kunnen doen. Je had naar buiten kunnen gaan, je had kunnen bepalen om niet te stemmen, je had een O kunnen tekenen in plaats van een X et cetera. De mogelijke redenen voor het uitvoeren van een handeling, in dit geval dat je voor Bush wilde stemmen, zijn niet causaal bepalend voor het uitvoeren van een bepaalde handeling (Searle 2004, 293-94). Searle claimt dat er een waarneembaar 'gat' bestaat tussen deze redenen voor het uitvoeren van een handeling en de handeling zelf. Dit gat wordt volgens Searle ook wel de vrijheid van de wil genoemd. In dit gat is het volgens Searle het zelf die besluiten maakt en de verschillende redenen evalueert. De verschillende redenen om een handeling uit te voeren (zoals de motivaties, overtuigingen en verlangens) zijn conflicterend en veronderstellen een zelf die deze redenen bewust evalueert om tot een besluit te komen (Searle 2001, 30-31).

Het maakt voor Searle niet uit of vrije wil in ontologische zin echt bestaat of dat het een illusie is, omdat het gat op geen enkele manier kan worden weggedacht. Zelfs als je zou geloven dat alles gedetermineerd is en besluit om geen handelingen meer te verrichten, dan kan het besluit om geen handelingen te verrichten alleen maar worden gezien als je eigen besluit als je vrije keuze veronderstelt. Dan heb je vrij gekozen om niet meer vrij te kunnen kiezen (Searle 2005, 15). Dit gevoel van vrijheid staat volgens Searle dus los van het vrije wil debat. De vrije keuze waar Searle over spreekt is niet afhankelijk van of de realiteit gedetermineerd is of niet. Hij stelt dat wij niet weten of vrije wil in ontologische



zin bestaat, maar stelt wel dat de psychologische vrijheid die we ervaren wel echt bestaat (Searle 2004, 164).

Om dit gevoel van vrije keuze te kunnen verklaren moeten we volgens Searle op zijn minst een zelf veronderstellen die functioneert als de 'locus van initiatie van acties'. Besluiten nemen en handelingen uitvoeren zijn niet zomaar acties die gebeuren, er is een *ik* die besluiten neemt en een *ik* die handelt. Deze 'ik' is echter niet het zelf waar Hume kritiek op leverde, maar een puur formele entiteit. Het is een formele entiteit, omdat het niet terug te vinden is in de percepties. Er is geen enkele ervaring van. (Searle 2005, 295).

Deze formele entiteit is dus in staat om acties te initiëren en uit te voeren. Searle claimt daarnaast dat deze entiteit ook dezelfde entiteit moet zijn als de entiteit die reflecteert op de redenen waarom het bepaalde acties heeft uitgevoerd. Ook moet het dezelfde entiteit zijn als de entiteit die een repertoire aan waarnemingen en herinneringen bezit die de basis vormen van de redenen waarom het handelt zoals het handelt. Als de entiteit, die handelt om een bepaalde reden, namelijk niet dezelfde entiteit is als de entiteit die waarneemt, herinnert en reflecteert dan zou je niet de benodigde connectie kunnen maken om de handelingen te kunnen verklaren (Searle 2005, 16). Als bijvoorbeeld een bus op je af komt gereden en je springt ervoor aan de kant, dan moet de entiteit die de beslissing neemt om te springen, dezelfde entiteit zijn als de entiteit die de bus waarnam. Daarnaast moet het ook dezelfde entiteit zijn als de entiteit die later als de bus is doorgereden reflecteert op de gemaakte handeling om dezelfde reden.

Searle beargumenteert dat het niet gek is dat veel mensen een homunculus postuleren die vanuit het brein alles waarneemt, besluiten neemt en hierop reflecteert. Wij ervaren de werkelijkheid namelijk ook zo. Om onze bewustzijnservaringen te begrijpen moeten we volgens Searle dus een zelf postuleren die bewust is, die door de tijd heen hetzelfde blijft, percepties en herinneringen heeft, die handelt vanuit redenering in het gat van de vrijheid van

de wil, die in het gat de mogelijkheid heeft om te beslissen en die op zijn minst verantwoordelijk is voor een deel van het vertoonde gedrag (Searle 2005, 16). Zoals eerder beschreven is deze zelf geen homunculus, zoals ego-theoretici claimen, maar een formeel principe van het bewustzijnsveld. We zouden een categoriefout maken als we het zelf los zouden zien van het bewustzijnsveld, zoals ego-theoretici dat wel doen. Dit zou dezelfde fout zijn als het aanwijzen van een paar gebouwen bij de vraag waar de universiteit zich bevindt. De universiteit is namelijk geen ding dat je kunt aanwijzen, maar een organisatie die gedefinieerd wordt door de interactie tussen de professoren, studenten, werknemers et cetera (Ryle 1949, 21-4). Zo is het zelf volgens Searle ook geen ding, maar een formeel of logisch principe dat bepaalt hoe het bewustzijnsveld zich structureert en organiseert (Searle 2005, 17).

We kunnen het formele zelf volgens Searle beter begrijpen als we het zien als een soort perspectief of gezichtspunt. Net zoals je alle visuele stimuli alleen maar kunt begrijpen als je ervanuit gaat dat ze allemaal worden waargenomen vanuit een bepaald gezichtspunt, zo kun je alle bewuste ervaringen ook alleen maar begrijpen als je ervan uitgaat dat ze allemaal worden waargenomen door een temporeel continue formele zelf. Deze zelf is de waarnemer en manipulator van bewuste ervaringen, maar is zelf niet waarneembaar en behoort niet tot de content van het bewustzijnsveld.

### **Searle's valkuilen**

Zoals ik heb laten zien claimt Searle dat er een zelf gepostuleerd moet worden om te begrijpen hoe wij de werkelijkheid ervaren. We kunnen zonder een zelf, volgens Searle, niet verklaren waarom onze waarneming van de werkelijkheid wordt gekenmerkt door eenheid. Er moet een zelf zijn die alle verschillende bewuste ervaringen met elkaar verbindt. Daarnaast kunnen we zonder een zelf te postuleren ook niet begrijpen waarom we het gevoel hebben dat we op elk moment van de tijd het gevoel hebben vrije keuzes te kunnen maken.

Searle begaat in zijn beide argumenten echter grove redeneerfouten. Ten eerste claimt hij dat het gevoel van de eenheid van ervaring voldoende bewijs is voor het bestaan van een formele zelf die alle ervaringen aan elkaar bindt. Er zijn echter ook alternatieve verklaringen mogelijk voor de gevoelde eenheid van ervaring die geen bestaan van een zelf postuleren. Searle geeft daarnaast ook geen verklaring voor hoe deze formele zelf dan deze ervaringen zou binden, dus een sterk argument is het niet. De formele zelf lost dus niks op in het filosofisch debat omtrent de eenheid van ervaring. Het creëert meer vragen dan antwoorden. Ten tweede claimt Searle dat er een formele zelf gepostuleerd moet worden om te kunnen begrijpen waarom het voor ons voelt dat we vrije keuzes maken. Ook dit argument is te ontkrachten door een alternatieve verklaring te bieden voor waarom het voor ons voelt dat we vrije keuzes maken, zonder een zelf te hoeven postuleren.

In de volgende hoofdstukken zal ik beide argumenten voor het bestaan van een formele zelf ontkrachten. Ik zal hierbij vooral gebruik maken van de HOG theorie in de traditie van Rosenthal. Door deze theorie toe te passen op de argumenten van Searle, voor het bestaan van een temporeel continue formele zelf, zal ik laten zien dat de HOG's alle functies kunnen vervullen die Searle toeschrijft aan een zelf, zonder een zelf te veronderstellen.

## **De HOG theorie**

De HOG theorie is een theorie die een verklaring kan geven voor waarom sommige mentale staten bewust zijn en sommige niet. Het beschrijft daarnaast ook de mechanismes die voor deze bewustwording zorgen. Volgens deze theorie zijn bewuste staten 'mentale staten die op een bepaalde manier in relatie staan met een mentale staat van een hogere orde' (Rosenthal 2002, 235-6). Volgens Rosenthal is het bewustzijn dat we van een mentale staat hebben niks anders dan

het hebben van een gedachte over datgene. Mentale staten zijn eerste-orde staten die op zichzelf niet bewust zijn. Hieronder vallen onder andere zintuigelijk input, herinneringen, overtuigingen en eerste-orde gedachtes. Mentale staten hebben om bewust te kunnen worden altijd een HOG nodig, volgens Rosenthal. Mentale staten worden dus alleen bewust als de gedachtes van een hogere orde deze mentale staten adequaat representeren. Je zou ook kunnen zeggen dat bij bewustwording deze eerste-orde mentale staten als content dienen voor een tweede-orde HOG.

Stel dat iemand bijvoorbeeld visuele input binnenkrijgt van een vliegtuig dat overvliegt. Volgens de HOG theorie wordt men alleen bewust van het vliegtuig als die op een bepaalde manier gerepresenteerd wordt door een HOG. Zonder de HOG zou de visuele input blijven hangen in de lagere orde mechanismes van het brein, zoals de visuele cortex, waardoor het onbewust zou blijven en men het vliegtuig niet bewust zou waarnemen. Rosenthal stelt dat deze HOG's vooral opereren in de breindelen waarvan bewezen is dat ze betrokken zijn bij hogere orde operaties, zoals de prefrontale en pariëtale cortex (Lau en Rosenthal 2011, 365).

Tweede-orde HOG's representeren dus eerste-orde mentale staten, maar ze manipuleren deze mentale staten niet. Het is niet zo dat een HOG intrinsieke verandering teweegbrengt in een mentale staat waardoor het bewust wordt. De HOG maakt de mentale staat bewust door er simpelweg te zijn. Bewustzijn is in de HOG theorie dus relationeel en niet een intrinsieke eigenschap van bewuste staten.

HOG's zijn zelf vaak onbewuste gedachtes die er dus voor kunnen zorgen dat mentale staten bewust worden als het de content op een adequate manier representeert. Wanneer de content op de juiste manier wordt gerepresenteerd door een HOG word je op een onmiddellijke wijze bewust van de content. De HOG wordt hierin echter niet zelf bewust waargenomen. HOG's zijn alleen maar potentiële bewust makers (Rosenthal 2002, 243). Toch kunnen de HOG's wel

bewust worden waargenomen als ze door een gedachte van een nog hogere orde worden gerepresenteerd. Dit kan volgens Rosenthal gedaan worden door middel van introspectie (Rosenthal 2002, 242). Je kan dus volgens Rosenthal spreken over zelfbewustzijn wanneer een tweede-orde HOG de content als content wordt genomen voor een derde-orde HOG.

### **De HOG theorie en de neurowetenschap**

De HOG theorie is een filosofisch concept dat aansluit bij de neurowetenschappelijke bevinding die zijn gedaan met betrekking tot bewuste en onbewuste processen. Searle praat over bewustzijn als een biologisch fenomeen dat ook op een biologisch manier verklaard zou moeten worden. Dit is een van de redenen waarom ik de HOG theorie zie als een goede kandidaat om het bestaan van een temporeel continue formele zelf, zoals Searle die postuleert, te ontcrachten. De HOG theorie maakt namelijk gebruik van de wetenschap die het bewustzijn, net zoals Searle, op neurobiologisch niveau probeert te verklaren.

Rosenthal beargumenteert dat de hogere-orde representaties verwerkt worden door de prefrontale cortex. De eerste-orde mentale staten zouden daarentegen tot het domein van de lage-orde breinmechanismen horen, zoals de eerder beschreven visuele cortex. Uit neurowetenschappelijke onderzoeken die zijn gedaan naar de functie van de prefrontale cortex blijkt inderdaad dat de prefrontale cortex voornamelijk actief is bij taken van hoge complexiteit zoals planning, cognitieve controle en het orkestreren van hogere-orde gedachtes (Miller en Cohen 2001, 167; Brown, Lau, en LeDoux 2019, 761-62). Ook laten neuro-imaging onderzoeken zien dat de prefrontale cortex betrokken is bij processen die veel bewuste aandacht eisen (Lau en Rosenthal 2011, 367-68). In een van deze neuro-imaging onderzoeken lieten Lau en Rosenthal participanten taken uitvoeren waar verschillende niveaus van bewuste aandacht voor nodig was. Tijdens deze taken werd de hersenactiviteit gemeten en werd gekeken naar welke breindelen het meest actief waren bij de taken die de meeste bewuste

aandacht vereisten. Hieruit bleek dat, van alle breindelen, de prefrontale cortex altijd het meest actief was bij taken waar veel bewuste aandacht voor nodig was.

Rosenthal beargumenteert daarnaast dat hogere-orde gedachtes, die geconstrueerd worden in hogere-orde breindelen, gebruik maken van lagere-orde gedachtes, die plaatvinden in lagere-orde breindelen. Dit zou betekenen dat onbewuste processen ontkoppeld zijn en apart functioneren van bewuste processen, maar wel als content kunnen dienen voor deze bewuste processen. Neuropsychologisch onderzoek laat zien dat onbewuste processen inderdaad kunnen doorwerken, zelfs wanneer de bewuste processen zijn aangetast. Zo blijkt dat schade aan de prefrontale cortex ervoor zorgt dat de mogelijkheid tot subjectieve verslaggeving wordt aangetast terwijl de taakprestatie op hetzelfde niveau blijft (Lau en Rosenthal 2011, 368-69).

Neuropsychologisch en -wetenschappelijke onderzoek laat dus inderdaad zien dat de functies van de prefrontale cortex voornamelijk bestaan uit complexe hogere-orde taken, zoals het orkestreren van gedachtes. Ook is bewezen met neuro-imaging onderzoeken dat de prefrontale cortex vooral actief is bij taken waar veel bewuste aandacht voor nodig is. Daarnaast blijken de onbewuste processen juist in de lagere-orde breinmechanismen plaats te vinden en laat onderzoek zien, met betrekking tot de relatie tussen onbewuste en bewuste processen, dat lagere-orde breinmechanismen kunnen blijven functioneren, zelfs wanneer de hogere-orde breinmechanismen zijn aangetast. De aannames van Rosenthal over de neurologische basis van de HOG's zijn dus niet vergezocht. Maar naast bewijs voor een neurologische basis van de HOG theorie zijn er ook een aantal objecties tegen de HOG theorie in te brengen.

### **Objecties tegen de HOG theorie**

De twee meest besproken objecties zijn de *targetless higher-order representation problem* (Neander 1998, 424-25; Block 2011, 21-2) en de *'rock' objection* (Goldman 1993, 366-67). Het eerste probleem stelt dat de HOG theorie niet kan

verklaren waarom we bewust kunnen worden van dingen die we niet hebben waargenomen. Zo kan bepaalde stimuli verkeerd worden geïnterpreteerd waardoor het voor de illusie van een waarneming zorgt. Dit zouden dan gedachtes van hogere orde moeten zijn zonder de representatie van een mentale staat van een lagere orde. Het tweede probleem beschrijft dat wanneer we bewust worden van een steen, dat de steen zelf niet bewust wordt. Waarom zou dit dan wel gelden voor gedachtes?

Deze objecties zijn echter beide overtuigend te weerleggen. Over de *targetless higher-order representation problem* zegt Rosenthal dat het inderdaad klopt dat we fenomenaal bewust kunnen zijn van HOG's zonder content, maar dat dit geen probleem hoeft te zijn voor de HOG theorie. Hij geeft een voorbeeld van mensen die bijvoorbeeld al pijn voelen bij de tandarts voordat er überhaupt een boor in de tand is gezet. De combinatie van angst inclusief het geluid van de boor laat de patiënten al pijn voelen, zonder dat er echte tactiele stimuli aan te pas komt. Dit is een voorbeeld waarin HOG's die gaan over het ervaren van pijn op zichzelf genoeg zijn om fenomenaal bewustzijn van pijn te induceren die niet verschilt van de pijn die ze daadwerkelijk zouden ervaren wanneer de boor er wel in zit (Rosenthal 2011, 436).

De *'rock' objection* van Goldman is weerlegt door Gennaro, een medevoorzitter van de HOG theorie. Hij beargumenteerde dat de HOG theorie niet veronderstelt dat zomaar alles als juiste content voor een HOG kan dienen. Rosenthal (2002, 243) stelde zelf, zoals eerder beschreven, ook dat HOG's alleen kunnen zorgen voor bewuste ervaringen als het de content op een adequate manier representeert. Een steen kan niet op een adequate manier gerepresenteerd worden, omdat het niet in een complexe mentale staat gecoöperend kan worden (Gennaro 2005, 21).

De HOG theorie staat dus in lijn met de bevindingen uit de neuropsychologie en- wetenschap en de hoofdobjecties die zijn gedaan tegen de HOG theorie zijn overtuigend te weerleggen. Daarnaast is de HOG theorie een

theorie die door verschillende filosofen en wetenschappers is bediscussieerd en verder is uitgewerkt (Rosenthal en Weisberg, 2008). Concluderend is de HOG theorie een standvastige theorie die gebruikt kan worden om Searle's idee over het bestaan van een temporeel continue formele zelf in twijfel te trekken. In de twee volgende hoofdstukken zal ik laten zien hoe deze HOG theorie kan worden gebruikt om de eigenschappen, die Searle toeschrijft aan de formele zelf, te verklaren zonder een temporeel continue entiteit te hoeven postuleren.

### **Kritiek op het zelf als de binder van ervaringen**

Het eerste argument dat Searle aanhaalt voor het bestaan van een zelf is een zwak argument dat niet logisch standhoudt om een aantal nader te bespreken redenen. Ten eerste hoeft de eenheid van ervaringen niet per se samengebonden te zijn door een en dezelfde formele entiteit. Het klopt dat wij de wereld fenomenologisch als een eenheid ervaren, maar dit betekent niet dat elke bewustzijnservaringen ook samengebonden wordt door dezelfde formele entiteit. Het kan zijn dat elk proces van binding een nieuwe 'binder' in het leven roept. Searle claimt bijvoorbeeld dat bewustzijn afwezig is in de slaap en pas weer actief wordt bij het opstaan. Wie zegt dat bij het opstaan dezelfde 'binder' aan het werk is ten opzichte van de dag ervoor? We kunnen hierbij Hume's scepticisme op het zelf ook toepassen op het formele zelf van Searle. Het beargumenteren dat er één formele entiteit bestaat die verantwoordelijk is voor het binden van ervaringen, omdat de werkelijkheid als een continue eenheid wordt ervaren, is hetzelfde als zeggen dat er maar één buschauffeur is die elke dag lijn 99 bestuurd, omdat de bus elke dag rijdt. Het klopt dat de werkelijkheid fenomenologisch op elk moment van de tijd ervaren wordt als een eenheid, net zoals het klopt dat lijn 99 elke dag rijdt, maar het proces van binding hoeft logisch gezien niet per definitie te betekenen dat de binding plaatsvindt door een en dezelfde entiteit, net zoals het



niet per se hoeft te betekenen dat de bus elke dag door dezelfde buschauffeur wordt bestuurd. Binding is een proces dat niet toegewijd hoeft te worden aan één formele zelf. De stelling dat de eenheid van ervaring veronderstelt dat er een formele entiteit moet bestaan die alle ervaringen aan elkaar bindt, houdt dus logisch gezien geen stand.

Daarnaast zou het ook mogelijk kunnen zijn dat het zelf juist datgene is dat voortkomt uit de binding van ervaringen en er niet aan vooraf gaat. Searle stelt dat de continuïteit van het bewustzijnsveld door de tijd heen wordt geclaimd door een zelf dat het bewustzijnsveld ziet als de continuïteit van zijn of haar eigen bewustzijn (Searle 2004, 292-93). Het zou dus ook kunnen zijn dat het gevoel, van het zijn van een temporeel continue zelf, voortkomt uit het claimen van het bindingsproces van ervaringen nadat het zelf er zelf uit is ontstaan. Het zelf wordt dan in dit geval dus gecreëerd door het bindingsproces en is niet zelf de binder van ervaringen. Dit laatste tegenargument zal ik nu verder uitwerken aan de hand van de HOG theorie.

### **De HOG theorie als verklaring voor de gevoelde eenheid van ervaring**

Searle claimt dat het gevoel, dat we de werkelijkheid ervaren als een eenheid, genoeg bewijs is voor het bestaan van een temporeel continue formele zelf. Hij geeft wel toe dat hij het gevoel dat we hebben, dat wie wij in essentie zijn gelijk staat aan deze temporeel continue zelf, niet kan verklaren.

'A second worry that I have is that I do not know how to account for the fact that an important feature of our experiences is what one might call a "sense of self"... My sense of self definitely exists, but it does not solve the problem of personal identity, and it does not yet so far flesh out the purely formal requirement that I said was necessary to supplement Hume's account in order to account for the possibility of free rational action. So, though this chapter is a beginning of a

discussion of the self, it is not more than a beginning' (Searle 2004, 298-99).

Ik beargumenteer dat de HOG theorie een goede alternatieve verklaring kan geven voor deze *sense of self* zonder een temporeel continue zelf te hoeven postuleren. Voordat ik hier echter iets over kan zeggen moet ik eerst laten zien dat de HOG theorie het gevoel, dat wij de werkelijk als een eenheid ervaren, ook kan verklaren.

HOG's opereren niet alleen, maar in clusters of bundels (Rosenthal 2003, 328). Een voorbeeld die goed illustreert hoe dit er in de praktijk uit ziet is het welbekende cocktail effect. Op een cocktailfeestje zijn er verschillende mensen tegelijk aan het praten. Als er geen aandacht wordt gevestigd op een bepaald gesprek klinkt al het geluid om je heen als ruis. Als iemand echter je naam roept spits je meteen de oren en hoor je wat diegene over je zegt, terwijl als diegene je naam had weggelaten je waarschijnlijk niks van het gesprek had meegekregen. Het in een keer bewust worden van je naam kan alleen mogelijk zijn als je de woorden van de verschillende gesprekken, die gaande waren op hetzelfde moment, gescheiden van elkaar hoorde. Je was je voordat je naam werd genoemd alleen maar bewust van ongedifferentieerd achtergrondgeluid. Vanuit het perspectief van de HOG theorie betekent dit dat de HOG, die ervoor zorgde dat alle geluiden los van elkaar werden gehoord, het horen van die geluiden representeerde als een enkelvoudig ongedifferentieerde groep van waarnemingen. Dit werkt niet alleen voor stimuli die binnenkomt via de oren, maar voor alle mogelijke modaliteiten waarop wij content voor de HOG's kunnen binnenkrijgen (Rosenthal 2003, 328-9).

De clusterwerking van de HOG's induceren dus een bewust gevoel van eenheid tussen de mentale staten. We zijn ons immers, zoals eerdergenoemd, op een onmiddellijke manier bewust van de content van HOG's en ze representeren de content in clusters en groepen. Er is echter geen ultieme HOG die alle bewuste staten omvat. Er is dus geen ultieme Gedachte, zoals James (1890, 401) die wel

voor de geest had. Maar hoe kan de HOG theorie dan verklaren dat we voelen alsof we de werkelijkheid als een eenheid ervaren, terwijl er verschillende HOG's tegelijk actief zijn die niet met elkaar verbonden zijn? En hoe kan dit het gevoel verklaren, waar Searle ook problemen mee heeft, dat we altijd het idee hebben dat we een temporeel continue zelf zijn die bewust is van onze gedachtes? Deze vragen zal ik in de volgende paragraaf beantwoorden.

### **De HOG theorie als alternatief voor een temporeel continue formele zelf**

Een HOG is een gedachte die ervoor zorgt dat iemand zich in een bepaalde mentale staat of in een cluster van mentale staten bevindt. Een HOG refereert niet alleen maar naar de mentale staat, maar ook naar zichzelf als diegene die de mentale staat heeft. In elke HOG schuilt een indexicale 'ik' (Caruso 2012, 225). Deze zelfreferentie is onontkoombaar. Een gedachte over iets hebben, maakt het ding alleen maar bewust als de gedachte het representeert als iets dat aanwezig is (Rosenthal 2005, 340). Bewust zijn van de mentale staat dat het ding representeert als aanwezig, maakt je er dus bewust van dat de mentale staat aan iemand toebehoort. Alle HOG's hebben daarom de content 'ik ben in mentale staat x' (Rosenthal 2005, 344). Alhoewel Rosenthal het niet expliciet zo noemt, zou je kunnen zeggen dat tweede-orde HOG's zorgen voor bewustzijn en derde-orde HOG's bewustzijn van het bewustzijn, wat gelijkstaat aan zelfbewustzijn.

Deze zelfreferentie is echter niet genoeg om te kunnen verklaren waarom we de werkelijkheid altijd ervaren alsof het wordt waargenomen door een en dezelfde temporeel continue zelf. Het zou namelijk net zo goed kunnen zijn dat elke HOG refereert naar een andere zelf. Het gevoel van temporele continuïteit kan dan alleen maar tot stand komen als alle HOG's naar dezelfde zelf refereren. Maar hoe is dit mogelijk als niet alle HOG's met elkaar zijn verbonden?

Elke HOG ziet het zelf waarna het refereert als de drager van de mentale staten en dus als de denker van de HOG zelf. Niks in de karakterisering hiervan impliceert dat het zelf in de ene HOG hetzelfde is als die in de andere HOG's.

Echter, er is ook geen verschil aan te wijzen tussen de verschillende dragers van mentale staten (Rosenthal 2005, 346). Aangezien, zoals eerder is beschreven, we door de HOG's op een onmiddellijke manier bewust worden van de content, lijkt het alsof alle ervaringen worden waargenomen door dezelfde temporeel continue zelf. Het formele zelf die Searle postuleert is in deze vergelijking dus niks meer dan een vluchtige derde-orde HOG. Deze derde-orde HOG's zijn echter niet temporeel continu en zijn contingent van aard.

Wanneer tweede-orde HOG's de content vormen voor een derde-orde HOG, dan vindt er, zoals eerdergenoemd, zelfbewustzijn plaats. Dit zelfbewustzijn geeft ons een bewust gevoel van het zijn van een temporeel continue zelf, terwijl elk moment van zelfbewustzijn eigenlijk een nieuwe *sense of self* in het leven roept. Introspectief bewustzijn zorgt er dus voor dat wij het gevoel krijgen dat alle bewuste staten met elkaar verbonden zijn in één enkelvoudige bewuste zelf, terwijl dit niet het geval is. Er is gewoon geen verschil aan te wijzen tussen de verschillende zelden die zijn ontstaan uit de bewustwordingen, waardoor het lijkt alsof er één enkelvoudige temporeel continue zelf bestaat.

Door de HOG theorie toe te passen op het probleem van het zelf kunnen we dus verklaren waarom wij ons een temporeel continue drager van mentale staten voelen. Het probleem is echter dat we ons niet alleen als een drager van mentale staten voelen, maar als iets veel complexer. Wij identificeren onszelf namelijk ook met persoonlijke markers zoals onze persoonlijke geschiedenis, lichaamseigenschappen, karaktereigenschappen et cetera. Hoe kan de HOG theorie dit verklaren?

Rosenthal stelt dat zelfidentificatie bestaat uit het identificeren wie het is waar je naar verwijst, wanneer je tegen jezelf praat of denkt (Rosenthal 2005, 348-49). Dit individu waar je naar verwijst, dat uit deze eerste-persoons gedachten wordt gepostuleerd, bestaat uit verschillende contingente persoonlijke markers. Elke nieuwe eerste-persoons gedachte heeft toegang tot de voorgaande eerste-persoonsgedachte om op voort te bouwen. Dit zorgt ervoor dat over de tijd heen

er een steeds groter repertoire aan zelf identificerende gedachtes beschikbaar komt. Dit repertoire bevindt zich in het geheugen. Hier had Searle dus wel gelijk in. Het geheugen is een belangrijk aspect voor ons eerste-persoonsperspectief en het gevoel van het zijn van een temporeel continue zelf. Dit gevoel is echter, zoals ik eerder heb beargumenteerd, een illusie.

Elke HOG representeert de mentale staat alsof het dus aan een bepaald individu toebehoort. Door nieuwe eerste-persoons gedachtes te hebben breiden we het repertoire van persoonlijke markers, waarmee we onszelf identificeren, uit. We identificeren daarom het individu waar alle HOG's naar verwijzen als één enkelvoudige temporeel continue zelf die dezelfde denker is van alle HOG's. Als we aan introspectie doen, worden wij bewust van deze HOG's en lijkt het voor ons alsof alle HOG's naar hetzelfde individu verwijzen. Hierdoor worden we bewust van onszelf als het middelpunt van bewustzijn (Rosenthal 2005, 351). Dit betekent echter niet dat het individu dat gepostuleerd wordt door de introspectieve HOG's daadwerkelijk bestaat, zelfs niet in formele vorm. Het zelf waarnaar in de ene HOG verwezen wordt verschilt met het zelf waarnaar in een andere HOG verwezen wordt.

Kort samengevat bestaat er dus geen temporeel continue formele zelf. Het gevoel dat we dit wel zijn is te verklaren door de HOG theorie. De HOG's bevatten allemaal een indexicale 'ik'. Je zou dit kunnen zien als een minimaal concept van een zelf. Het lijkt of deze zelf in elke HOG dezelfde zelf is, omdat er geen verschil in zit. Wanneer tweede-orde HOG's gerepresenteerd worden door derde-orde HOG's, ervaren wij een gevoel van zelfbewustzijn. We zijn ons dan bewust als de drager van mentale staten. Wanneer we naar onszelf verwijzen, maken we gebruik van persoonlijke markers die liggen opgeslagen in het geheugen. Een tweede-orde HOG kan deze persoonlijke markers representeren om er bewust van te worden en hiermee een individu te creëren waarmee het zelf, die op dat moment ontstaan is, zich kan identificeren. Dit zorgt voor een gevoel van temporeel continuïteit. Dit repertoire aan persoonlijke markers is echter contingent en kan per situatie

verschillen. Zo kun je in verschillende situaties een andere persoonlijkheid hebben, maar je toch altijd hetzelfde individu voelen. Dit komt doordat alle HOG's de content 'ik ben in mentale staat X' hebben. Het maakt dus niet uit welke persoonlijke markers er worden gerepresenteerd, je hebt in dat opzicht in elke situatie het gevoel dat je jezelf bent.

De HOG theorie kan dus verklaren waarom wij de werkelijkheid als een eenheid ervaren en waarom het lijkt alsof die eenheid wordt waargenomen en gecreëerd door een temporeel continue zelf, zonder een temporeel continue formele zelf te hoeven postuleren. De HOG theorie biedt een goede alternatieve verklaring en kan worden gebruikt om te laten zien waarom het niet nodig is om een formele zelf te postuleren om dit gevoel van eenheid te kunnen verklaren. Door de HOG theorie toe te passen op de theorie van Searle kunnen we het formele zelf ontmaskeren als een onnodige gepostuleerde entiteit.

### **Kritiek op het zelf als de maker van keuzes**

Het tweede argument dat Searle aanhaalt, die stelt dat er een formele zelf gepostuleerd moet worden om te begrijpen waarom het voor ons voelt dat we vrije keuzes maken, blijkt ook een argument te zijn dat niet moeilijk te ontcrachten is. Searle claimt dat er een gat bestaat tussen onze mogelijke redenen voor handelen en de handeling zelf. Ik zal laten zien dat dit gat een illusie is dat in werkelijkheid niet bestaat als we kijken naar recente bevindingen in de psychologie en de neurowetenschap.

Searle stelt daarnaast dat dit gat niet weggedacht en gevoelt kan worden en dat het hoe dan ook een onderdeel zou blijven van onze fenomenologische ervaring van de werkelijkheid. Ik zal laten zien dat dit gevoel van psychologische vrijheid ook verklaard kan worden door de HOG theorie, zonder een formele zelf in het leven te roepen die zogenaamd de psychologische touwtjes in handen heeft.

### **Searle's illusionaire gat van psychologische vrijheid**

Searle stelt dat er een gat zit tussen de redenen voor ons handelen en de handeling zelf. Er zijn verschillende redenen voor gedrag aan te wijzen die niet causaal bepalend zijn voor het gedrag dat uiteindelijk uitgevoerd wordt. Volgens Searle is het het formele zelf die in dit gat opereert en uit de verschillende redenen voor handelen een keuze maakt. 'Nothing fills the gap, you just make up your mind to do something' (Searle 2001, 17). Het is hierin, volgens Searle, het formele zelf die de laatste beslissing maakt en dus de causale macht in handen heeft (Searle 2001, 94-5). Volgens Searle komt er input binnen in het brein, zowel extern door de zintuigen, als intern door bijvoorbeeld bestaande overtuigingen. Hierna maakt het zelf een afweging, neemt een beslissing en komt er output uit in de vorm van een handeling. Deze redenering gaat er dus vanuit dat het formele zelf een ingekapselde beslissingsmaker is die opereert in het gat van psychologische vrijheid (Searle 2004, 217).

Searle claimt dat iemand die gegijzeld wordt niet vrij is om te handelen zoals hij zelf wil en dat daardoor zijn psychologische vrijheid wordt afgenomen. Dit is een voorbeeld waarin iemand fysiek wordt tegengehouden om bepaalde acties uit te voeren waardoor de mogelijkheden voor handelen in grote mate wordt beperkt. Searle geeft echter ook voorbeelden van mensen waarvan mentaal de psychologische vrijheden worden afgenomen. Een voorbeeld hiervan is een drugsverslaafde die per se een hit nodig heeft en daarom al zijn acties om die verlangens heen baseert. Een ander voorbeeld dat hij geeft is van iemand onder hypnose. Diegene heeft wel het idee vrij te kunnen handelen tijdens de hypnose, maar in het echt wordt diegene gestuurd door de hypnotiseur om bepaalde handelingen uit te voeren. Deze persoon heeft dus wel het gevoel van psychologische vrijheid, maar had het eigenlijk niet. Het gat van psychologische vrijheid is in dit voorbeeld dus een illusie. Searle zegt echter dat ons handelen in het dagelijks leven niet te vergelijken is met hoe een drugsverslaafde of iemand onder hypnose tot een handeling komt. De psychologische oorzaken van het

handelen zijn te verschillend (Searle 2004, 224-5). In het voorbeeld van hypnose heeft de gehypnotiseerde wel het gevoel van psychologische vrijheid, maar is hij niet bewust van zijn echte motivaties, overtuigingen en verlangens. Als diegene dit wel zou zijn, dan zou hij anders gehandeld hebben. Echte psychologische vrijheid veronderstelt, volgens Searle, dat men bewust is van de eigen motivaties (Searle 2004,225).

Wat Searle hier echter over het hoofd ziet, is dat wij in ons dagelijks leven ook vaak niet bewust zijn van onze motivaties, overtuigingen en verlangens. Searle claimt dat het gat van psychologische vrijheid een illusie is in het geval van hypnose, omdat diegene zich niet bewust is van de eigen motivaties. Dit gat is, volgens Searle, wel volledig operationeel bij het maken van alledaagse keuzes waarin we wel vrij kunnen handelen. Searle gaat echter niet verder in op hoe bewust wij daadwerkelijk zijn van onze eigen motivaties, overtuigingen en verlangens in het dagelijks leven, terwijl dit wel een cruciaal punt is waarop hij het bestaan van een gat van psychologische vrijheid baseert.

Ik beargumenteer dat wij in het dagelijks leven minder bewust zijn van onze motivaties dan Searle doet vermoeden. Wij worden in het dagelijks leven namelijk ook door vele externe invloeden, waar we niet bewust van zijn, gestuurd. Daarnaast zijn we ons net als de gehypnotiseerde niet altijd bewust van onze eigen motivaties. Zo zijn bijvoorbeeld winkelketens zo ingericht dat de producten die ze het meest verkocht willen hebben op ooghoogte staan en dat de duurdere producten hoger op de schappen staan om mensen een gevoel te geven dat ze daadwerkelijk een kwalitatief beter product kopen. We worden daarnaast gebombardeerd met aanbiedingen die grotendeels onbewust worden waargenomen, maar alsnog ons gedrag beïnvloeden.

Een ander voorbeeld is dat van priming. Priming is het al dan niet onbewust aanbieden van stimuli, waardoor soortgelijke informatie sneller wordt herkend en waarop sneller wordt gereageerd (Molden 2014, 1-2). In de psychologie is priming een middel om het gedrag van participanten onbewust te beïnvloeden. De prime



wordt onbewust aangeboden, waarna participanten bepaalde taken moeten voltooien. De participanten handelen significant boven kans niveau alsof ze de prime bewust hebben waargenomen. Wanneer echter wordt gevraagd waarom ze een bepaalde handelingen hebben uitgevoerd, benoemen ze de prime niet als oorzaak, maar altijd de eigen motivaties. Ze kunnen de prime ook niet als oorzaak noemen, want ze waren er niet bewust van. Dit is een voorbeeld waarin het gedrag van mensen onbewust wordt beïnvloed, en daardoor ook hun psychologische vrijheid wordt afgenomen, terwijl ze er niet bewust van zijn. Dit is vergelijkbaar met hoe een gehyponotiseerde handelt door de instructies van een hypnotiseur. Diegene onder hypnose zal ook nadat het uit hypnose komt een logische verklaring bedenken voor het vertoonde gedrag.

Er is genoeg bewijs in de psychologie te vinden dat deze priming ook doorwerkt in het alledaags leven. Continu om ons heen worden we geprikkeld door stimuli waar we ons niet bewust van zijn. Onbewuste priming met voedselreclames, zorgden er bijvoorbeeld voor dat mensen meer gingen eten, terwijl ze bewust aangaven niet meer honger te hebben dan daarvoor (Harris, Bargh en Brownell 2009, 409-10). Dit effect werd ook gevonden bij het zien van een reclamebord met eten dat voor een aantal milliseconden vertoond werd tijdens het auto rijden. Bij de vraag wat voor avondeten de mensen zouden willen eten werd het eten dat onbewust werd waargenomen op het verkeersbord significant vaker gekozen dan je op kans niveau zou verwachten.

Als de mensen uit deze voorbeelden worden gevraagd waarom ze bepaald keuzes hebben gemaakt schrijven ze deze altijd toe aan eigen motivaties, terwijl dit dus feitelijk niet zo is. Zo zijn er nog vele andere invloeden die ons gedrag bepalen zonder dat we daar bewust van zijn. Zo worden bijvoorbeeld onze eigen motivaties en overtuigingen beïnvloed door de omgeving waarin we ons bevinden (Kaufmann 2002, 470-71). Je zou dus als tegenargument kunnen geven, tegen het bestaan van een gat tussen de mogelijke redenen voor handelen en de handeling zelf, dat we simpelweg niet bewust zijn van alle mentale staten die aan een

handeling voorafgaan. Deze kloof van onbewuste mentale staten zorgt op bewust niveau voor de illusie van een gat waarin we psychologische vrijheid ervaren.

Searle stelt dat we de indruk hebben dat we op elk moment van de tijd zelf kunnen bepalen wat voor acties we gaan ondernemen (Searle 2005, 15). We hebben volgens hem altijd een bepaald aantal mogelijkheden waaruit we er een kunnen kiezen om op te handelen. Uit het hierboven genoemde voorbeeld blijkt echter dat deze vrijheid door vele externe onbewuste oorzaken kunnen worden ingeperkt. We zijn ons er bijvoorbeeld niet van bewust dat we in het dagelijks leven continu geprimed worden. Zo kunnen we dit argument ook doortrekken en zeggen dat er vele andere mentale staten waar we ons ook niet bewust zijn. Dit hoeven niet per definitie mentale staten te zijn die veroorzaakt zijn door externe stimuli, maar kunnen ook interne staten zijn zoals impliciete attitudes of onbewuste overtuigingen. Hiermee kunnen we Searle's claim, die zegt dat de redenen voor het handelen niet causaal bepalend zijn voor het uitvoeren van een handeling, onderuithalen. Deze redenen voor het uitvoeren van een handeling, zoals de motivaties, overtuigingen en verlangens, zijn dus wel causaal bepalend. Het is alleen dat we ons niet bewust zijn van het causale proces dat aan de handeling voorafgaat. Het gat dat gecreëerd wordt door onze onwetendheid is dus een illusie zonder ontologische basis.

Searle zegt dat iemand die gegijzeld wordt niet psychologische vrij is, omdat zijn mogelijkheid om keuzes te maken flink wordt ingeperkt. Men zou dus hetzelfde kunnen beargumenteren voor Searle's formele zelf. Deze zelf wordt net zo goed ingeperkt door alle onbewuste invloeden die het gedrag van het individu bepalen. Daarnaast kan het ook zo zijn dat we ons niet bewust zijn van de causale keten van mentale staten die aan een handeling voorafgaan. Of men kan spreken van het bestaan een gat tussen de mogelijke redenen voor handelen en de handeling zelf is dus nog maar de vraag. Psychologisch gezien wordt dit gat dan gedicht door de verschillende externe en interne invloeden waar we ons niet bewust van zijn.

Zelfs als het gat niet bestaat en Searle hier een ontologische fout maakt, kan hij alsnog beargumenteren dat wij dit gat wel ervaren. Het punt dat hij namelijk wil maken is dat wij altijd het gevoel hebben alsof we een zelf zijn die uit bepaalde mogelijkheden voor handelen kan kiezen. Dit gat is misschien logisch wel weg te redeneren, maar kan fenomenologisch niet weg worden ervaren. We blijven altijd het gevoel hebben dat wij, als zelf zijnde, op elk moment in de tijd, zelf kunnen bepalen welk van de mogelijke redenen voor handelen wij kiezen. Er moet volgens Searle daarom een verklaring worden gegeven waarom we deze psychologische vrijheid ervaren. Zoals we eerder hebben kunnen zien is dit een van de redenen waarom Searle beargumenteert dat er wel een formele zelf moet bestaan. Ik beargumenteer echter dat de ervaring van psychologische vrijheid ook kan worden verklaard door hier de HOG theorie toe te passen. Hiermee kan het gevoel van psychologische vrijheid worden verklaard zonder dat er een temporeel continue formele zelf gepostuleerd hoeft te worden.

### **Het gevoel van psychologische vrijheid verklaard**

Mentale staten kunnen, zoals eerder beschreven, worden gebruikt als persoonlijke markers om jezelf als individu te identificeren. Deze mentale staten zijn bijvoorbeeld voorkeuren, overtuigingen en verlangens. Dit zijn eerste-orde staten die zijn gebaseerd op het geheugen. Deze persoonlijke markers waar we onszelf mee identificeren hebben echter weinig systematische connectie met elkaar en zijn contingente eigenschappen die ook per situatie kunnen verschillen (Rosenthal 2005, 348). Wanneer je een nieuwe keuze maakt, zonder dat er markers uit het verleden worden opgeroepen dan worden er nieuwe persoonlijke markers aangemaakt die later kunnen worden gebruikt in nieuwe HOG's (Rosenthal 2005, 349). De set van persoonlijke markers blijft dus continu groeien.

Wanneer je een bepaalde keuze maakt, bijvoorbeeld om naar een ijssalon te gaan om een ijsje te gaan eten, dan is de ervaring van die keuze een HOG die deze verschillende persoonlijke markers representeert samen met de mogelijke

keuze opties. Bij het maken van een keuze hoeft de HOG alleen maar de persoonlijke markers en opties te representeren die relevant zijn voor het maken van die keuze. Bij het voorbeeld van het besluiten om een ijsje te gaan eten, word je dus alleen maar bewust van de mogelijke redenen voor handelen, zoals Searle het noemt, die relevant zijn voor de afweging om een ijsje te gaan eten of niet. Dit zijn bijvoorbeeld de temperatuur buiten, hoeveel zin je hebt in een ijsje, of er iemand met je mee wilt gaan, of ze het ijs hebben dat je graag wilt hebben et cetera.

De representatie van de keuze bevat ook het concept van een zelf die de keuze maakt als het een derde-orde HOG betreft. Je bent je dan bewust als het zelf die bewust is van de verschillende persoonlijke markers waar het zich mee identificeert, inclusief de mogelijke keuze opties. Wanneer een keuze wordt gemaakt, dan representeert de HOG de keuze niet als gewoon een andere mentale staat, maar als een zelf die zich in die mentale staat bevindt en diegene is die de keuze maakt. Alle derde-orde HOG's refereren namelijk naar zichzelf als de drager van de mentale staten. Alle keuzes die worden gemaakt, waarvan men bewust is, roepen daarom automatisch een maker van keuzes in het leven. Hier moet wel benadrukt worden dat het gevoel van het zijn van een zelf die keuzes maakt, zoals eerder beschreven, niks meer is dan een vluchtige derde-orde HOG zonder enige ontologische spatio-temporele continuïteit (Rosenthal 2005, 364). Er is dus geen zelf die psychologische vrijheid ervaart, het is alleen het gevoel van het zijn van een zelf die psychologische vrijheid ervaart die bij het maken van een keuze tot stand komt.

De reden waarom wij deze keuze ervaren als een keuze die wij zelf hebben gemaakt en niet als een keuze die voor ons wordt gemaakt door voorgaande mentale staten is omdat de HOG's, zoals eerder beschreven, alleen maar mentale staten representeren die relevant zijn voor het maken van de keuze. Dit zorgt ervoor dat we ons niet bewust zijn van alle mentale staten die aan een keuze voorafgaan. Wanneer er uiteindelijk een beslissing wordt genomen, wordt die

beslissing ervaren als een niet-veroorzaakte oorzaak. Het lijkt dan alsof die beslissing ex nihilo wordt genomen door het zelf die tot stand komt door de representatie van een derde-orde HOG. Dit is ook de reden waarom wij na het maken van een keuze vaak het gevoel hebben dat we ook een andere keuze hadden kunnen maken. We waren ons niet bewust van alle mentale oorzaken die aan het maken van een keuze voorafgingen, omdat ze niet allemaal gerepresenteerd werden door de HOG in kwestie.

Het zou zelfs mogelijk kunnen zijn dat de beslissing al is gemaakt, zelfs voordat de persoonlijke markers en mogelijke opties überhaupt in het bewustzijn opkomen. Vooral als de keuze snel moet worden gemaakt, waardoor er geen tijd is voor derde-orde HOG's om het keuzemoment op een hogere orde te analyseren (Rosenthal 2005, 364). Een voorbeeld hiervan is bijvoorbeeld een professioneel voetballer die op de milliseconden nauwkeurig keuzes moet maken om uiteindelijk een doelpunt te kunnen scoren. Dit gebeurt allemaal automatisch en onbewust, maar wanneer later tijdens een interview wordt gevraagd hoe de voetballer tot zijn keuzes is gekomen kan diegene prima bewust reflecteren op de processen die voorafgingen aan de keuze. Daarnaast zal diegene ook altijd zichzelf zien als de maker van de keuzes, ook al gebeurden ze onbewust.

Een derde orde HOG kan in mijn interpretatie dus twee functies vervullen. Het kan op een hogere orde complexere informatie analyseren om een keuze te maken. Dit zou rationeel bewust overwegingen maken genoemd kunnen worden. De tweede functie is achteraf gemaakte keuzes analyseren, waardoor men hiervan kan leren. Bij beide functies heb je dus altijd het gevoel alsof jij de maker was van de keuze, zelfs als die onbewust gebeurde. Dit komt doordat, zoals eerder beschreven, een derde-orde HOG altijd het gevoel van het zijn van een zelf die de drager van die mentale staat is, omvat. Of de keuze onbewust is gemaakt, voordat het überhaupt in het bewustzijn opkomt, maakt dus geen verschil voor het gevoel van het zijn van een maker van keuzes.

Als een HOG alle mentale staten zou kunnen representeren dan zouden wij het gat van psychologische vrijheid, die Searle postuleert, ook niet ervaren, omdat er dan geen ruimte meer is om enige vrijheid te claimen. De HOG theorie kan dus worden toegepast om te laten zien waarom wij onszelf ervaren als een spatio-temporeel continue zelf die eigen psychologische vrijheid ervaart, zonder een zelf te hoeven postuleren om het argument sluitend te krijgen.

### **Initiëren van acties, reflectievermogen en het repertoire van herinneringen**

Searle schrijft aan zijn gepostuleerde formele zelf verschillende eigenschappen toe die het zou moeten bezitten om de ervaring van ons bewustzijnsveld te kunnen begrijpen. Het zelf moet in staat zijn om acties te initiëren, het moet kunnen reflecteren op de redenen waarom het bepaalde acties uitvoert en het moet toegang hebben tot een repertoire aan opgeslagen waarnemingen en herinneringen. Zonder een zelf die deze eigenschappen bezit zou je volgens Searle niet de benodigde connectie kunnen maken om handelingen te kunnen verklaren. Ik beargumenteer dat al deze eigenschappen ook aan de HOG's toegeschreven kunnen worden, zonder dat er een zelf gepostuleerd hoeft te worden.

Ten eerste zijn de HOG's in staat om acties te initiëren. Zoals ik eerder al heb beschreven kunnen derde-orde HOG's meerde functies bekleden. Een daarvan is het rationeel bewust analyseren van informatie. Wanneer deze informatie geanalyseerd is kan de content worden gebruikt voor andere mentale processen die uiteindelijk de handelingen in gang zetten. De HOG's kunnen dus functioneren als een schakel in de causale keten van het uitvoeren van een actie. Ten tweede zijn HOG's ook in staat om te reflecteren op de acties. Dit is de tweede functie van de derde-orde HOG's. Ze kunnen gemaakte keuzes representeren en deze hierdoor bewust maken alsof er een zelf is die de keuzes gemaakt heeft. Deze informatie zou kunnen worden opgeslagen in het repertoire van waarnemingen en herinneringen. HOG's hebben, ten derde, dus ook toegang tot het geheugen. Zoals eerder beschreven maken HOG's gebruik van persoonlijk markers om

zichzelf te identificeren. Deze persoonlijke markers bevinden zich in het geheugen en vormen een deel van het repertoire van alle waarnemingen en herinneringen. Je zou dus kunnen zeggen dat de HOG's alle functies van het formele zelf van Searle kunnen vervullen, zonder dat er een zelf als middelpunt van alle ervaringen gepostuleerd hoeft te worden.

## **Conclusie**

In deze thesis heb ik mij gebogen over het formele zelf waarvan Searle het bestaan postuleert. Ik begon met het kort omschrijven van de geschiedenis van het debat omtrent persoonlijke identiteit en het zelf. Hierin beschreef ik de twee uiteenlopende stromingen, die gecategoriseerd werden door Parfit als de ego- en de bundeltheorie van persoonlijke identiteit. Na deze korte inleiding zette ik Searle's positie in het debat uiteen. We zagen dat Searle zich voornamelijk bezighoudt met de vraag of er naast de sequentie van percepties een zelf gepostuleerd moet worden om onze ervaring van de werkelijkheid te kunnen begrijpen. Searle claimt contra-Hume dat dit wel het geval is. Deze zelf die Searle postuleert is echter geen zelf die te vinden is in percepties, maar een puur formele entiteit.

Het eerste argument dat Searle aanhaalt voor het bestaan van een formele zelf is de gevoelde eenheid van het bewustzijnsveld. Searle beargumenteert dat, wanneer er geen zelf wordt gepostuleerd, we niet kunnen verklaren waarom het bewustzijnsveld altijd gekenmerkt wordt door eenheid. Het tweede argument, dat Searle aanhaalt voor het bestaan van een zelf gaat over het gevoel van psychologische vrijheid. Searle claimt dat bij het maken van vrijwillige acties we altijd het gevoel hebben vrij te kunnen handelen. Daarnaast stelt hij dat de mogelijke redenen voor het uitvoeren van een handeling niet causaal bepalend zijn voor het daadwerkelijk uitvoeren van een bepaalde handeling. Er bestaat,

volgens Searle, een gat tussen de redenen voor het uitvoeren van een handeling en de handeling zelf.

Na deze uiteenzetting heb ik laten zien dat beide argumenten te ontcrachten zijn door de HOG theorie toe te passen op het probleem van het zelf, zoals Searle die omschrijft. Ik heb laten zien dat het eerste argument van Searle, dat de eenheid van ervaring bewijs is voor het bestaan van een binder van ervaringen, geen sterk argument is. Het gevoel van het zijn van een temporeel continue zelf, die de verschillende bewustzijnservaringen aan elkaar bindt, is ook op een zuinigere manier te verklaren. De HOG theorie beargumenteert namelijk dat in elke HOG een indexicale 'ik' of drager van mentale staten schuilt. Ik heb beschreven hoe HOG's mentale staten gegroepeerd representeren. Ook heb ik laten zien dat wanneer een derde-orde HOG een tweede-orde HOG representeert de tweede-orde HOG bewust wordt als de drager van de mentale staten die samen als een eenheid worden ervaren. Dit zorgt voor het gevoel alsof er een zelf is die in het middelpunt van de bewuste ervaring staat. Deze zelf is echter een illusie die bij elke derde-orde HOG tot leven wordt gebracht. Het zelf die in de ene HOG ervaren wordt, is niet dezelfde zelf als die in een andere HOG ervaren wordt. Er is dus geen temporeel continue zelf.

Het tweede argument dat Searle aanhaalt voor het bestaan van een formele zelf, dat wij zonder een zelf niet kunnen verklaren waarom we een gat ervaren tussen de mogelijke redenen voor handelen en de handeling zelf, heb ik in deze thesis ook ontcracht. Ik heb laten zien dat het gat waar Searle het over heeft niks meer is dan een illusie dat voorkomt uit ons niet-bewustzijn van onze eigen mentale staten. Ik heb ook laten zien dat, voor het gevoel van het bestaan van dit gat, een verklaring te geven is door de HOG theorie op dit probleem toe te passen. De HOG theorie stelt namelijk dat HOG's op een keuzemoment alleen de relevante persoonlijke markers en mogelijk opties bewust maakt. Wanneer een derde-orde HOG een tweede-orde HOG representeert tijdens het maken van een keuze wordt men bewust van zichzelf als de maker van keuzes. Dit gevoel van het



zijn van een zelf die keuzes maakt is echter niks meer dan een vluchtige derde-orde HOG zonder enige ontologische spatio-temporele continuïteit.

Ik heb in deze thesis dus laten zien dat het formele zelf, die Searle in het leven roept om bepaalde fenomenologische problemen op te lossen, een onnodig gepostuleerde entiteit is. De problemen van Searle omtrent persoonlijke identiteit en het zelf zijn op een zuinigere manier op te lossen, zonder het hoeven postuleren van entiteit waarvan we het bestaan niet kunnen bewijzen. Het wordt dus wel eens tijd om dat sterke gevoel, van het zijn van een zelf dat in het middelpunt van de ervaring staat, los te laten. We zijn nu al eeuwen lang op zoek naar het zelf, maar een stap dichterbij bewijs voor zijn bestaan komen we niet. Searle claimt met zijn theorie het oude lichaam-geest probleem te kunnen omzeilen, maar ik ben van mening dat hij, al dan niet met een omweg, weer terug bij af is. De geest is, na een rondje om in Searle's theorie, weer terug in de machine geplaatst.



## Referenties

- Ayer, Alfred. 1973. *The Central Questions of Philosophy*. London: Weidenfeld and Nicolson.
- Blackmore, Susan. 2013. "The Self." In *Consciousness: an Introduction*. New York: Routledge.
- Block, Ned. 2011. "The higher-order approach to consciousness is defunct" *Analysis* 71(3), 419–431.
- Bloom, Paul. 2004. *Descartes' Baby: How Child Development Explains What Makes us Human*. London: Heinemann.
- Brown, Richard, Hakwan Lau, en Joseph LeDoux. 2019. "Understanding the Higher-Order Approach to Consciousness." *Trends in Cognitive Sciences* 23(9): 754-768
- Caruso, Gregg. 2012. *Free Will and Consciousness: A Determinist Account of the Illusion of Free Will*. Lanham: Lexington books.
- Chalmers, David. 1996. *The Conscious Mind*. Oxford: Oxford University Press.
- Dennett, Daniel. 2003. "The Self as a Responding and Responsible Artifact." *Annals of the New York Academy of Sciences* 1001(1): 39–50.
- Descartes, René. 1972. *Treatise of Man*. Vertaald door Thomas Hall. Cambridge: Harvard University Press.
- . 2018. *Discourse on the Method: The Original Text with English Translation*. Vertaald en bewerkt door Constatin Vaughn. Glasgow: Erebus Society.
- Genarro, Rocco. 2005. "The HOT theory of consciousness: between a rock and a hard place." *Journal of Consciousness Studies* 12(2): 3–21.
- Goldman, Alvin. 1993. "Consciousness, Folk Psychology and Cognitive Science." *Consciousness and Cognition* 2(4): 264-82.
- Harris, Jennifer, John Bargh, en Kelly Brownell. 2009. "Priming Effects of Television Food Advertising on Eating Behavior." *Health Psychology* 28(4): 404–13.

- Hume, David. 1978. *Treatise of Human Nature*. Bewerkt door Lewis Selby-Bigge. Oxford: Clarendon Press.
- James, William. 1890. *The Principles of Psychology*. London: MacMillian.
- . 1904. "Does 'Consciousness' Exist?" *Journal of Philosophy, Psychology, and Scientific Methods* 1(1): 477-91.
- Kaufmann, Laurence. 2005. "Self-In-A-Vat: On John Searle's Ontology of Reasons for Acting." *Philosophy of the Social Sciences* 35(4): 447–79.
- Lau, Hakwan, en David Rosenthal. 2011. "Empirical Support for Higher-Order Theories of Conscious Awareness." *Trends in Cognitive Sciences* 15(8): 365-73.
- Locke, John. 1975. *An Essay Concerning Human Understanding*. Vertaald en bewerkt door Peter Nidditch. Oxford: Clarendon.
- Miller, Earl, en Jonathan Cohen. 2001. "An Integrative Theory of Prefrontal Cortex Function." *Annual Review of Neuroscience* 24(3): 167-202.
- Molden, Daniel. 2014. "Understanding Priming Effects in Social Psychology: What is 'Social Priming' and How does it Occur?" *Social Cognition* 32(3): 1-11.
- Nagel, Thomas. 1974. "What is it Like to be a Bat?" *Philosophical Review* 83(4): 435–50.
- Neander, Karen. 1998. "The Division of Phenomenal Labor: a Problem for Representational Theories of Consciousness." in *Language, Mind, and Ontology*. Oxford: Blackwell.
- Nimbalkar, Namita. 2011. "John Lock on Personal Identity." *Mens sana monographs* 9(1): 268-75. doi:10.4103/0973-1229.77443
- Parfit, Derek. 1987. "Divided Minds and the Nature of Persons." *Mindwaves*. Oxford: Blackwell.
- Putnam, Hilary. 1975. "The meaning of meaning In Mind, Language and Reality." *Philosophical Papers* 2(1): 215-271.
- Rosenthal, David. 2002. "Consciousness and the Mind." *The Jerusalem Philosophical Quarterly* 51(1): 227-251.

- . 2003. "Unity of Consciousness and the Self." *Proceedings of the Aristotelian Society* 103(1): 325-52.
- . 2005. *Consciousness and mind*. Oxford: Oxford University Press.
- , en Josh Weisberg. 2008. "Higher-order theories of consciousness." *Scholarpedia* 3(5): 4407.
- . 2011. "Exaggerated reports: reply to Block." *Analysis* 71(3): 431–437.
- Russell, Bertrand. 1966. *History of Western Philosophy*. London: Routledge.
- Ryle, Gilbert. 1949. *The Concept of Mind*. Chicago: University of Chicago Press.
- Searle, John R. 1992. *The Rediscovery of the Mind*. Cambridge: The MIT Press.
- . 2000. "Consciousness." *Annual Review of Neuroscience* 23(1): 557-78.
- . 2001. *Rationality in action*. Cambridge: MIT Press.
- . 2004. *Mind: A Brief Introduction*. Oxford: Oxford University Press.
- . 2005. "The Self as a Problem in Philosophy and Neurobiology." in *The Lost Self: Pathologies of the Brain and Identity*, bewerkt door Todd Feinberg en Julian Paul Keenen. Oxford: Oxford University Press.
- . 2016. "How Do Persons Maintain Their Identity?" Interview door Robert L. Kuhn. *Closer to Truth*, 6 september 2016.
- Wittgenstein, Ludwig. 1992. *Tractatus Logico-philosophicus*. Vertaald door David Pears en Brian McGuinness, London: Humanities Press International.